DOCUMENT RESUME

ED 144 739 RC 010 087

TITLE A Study of the Validity of Teacher Evaluations of Students in the Region 14 Education Service Center Cooperative for Migrant Education.
INSTITUTION Education Service Center Region 14, Abilene, Tex.
PUB DATE Mar 76
NOTE 7p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Criterion Referenced Tests; Elementary Education; *Evaluation Methods; *Migrant Education; *Student Evaluation; *Validity
IDENTIFIERS *Texas

ABSTRACT

The migrant program's evaluation system necessitates that instructional personnel evaluate students on a continuing basis using criterion-referenced items that have been developed and validated. This measurement system consists of 230 criterion-referenced items designed to measure 230 objectives in reading and math. These range across seven levels and grades K-6. The study determined whether teachers in the program were able to use the project developed criterion-referenced measurement system to produce valid measures on students. During February 1976, a stratified random sample of students was selected and matched with a group of randomly selected criterion-referenced items from the system. The project staff who developed the system (and on whom inter-rater reliability had been established) visited each school and individually tested the selected students on the selected items. Two measures were noted: the existing teacher rating (taken from student records) and the results of the staff testing. In both cases, teacher and staff ratings were in terms of whether the student did or did not perform satisfactorily on the objective. The resulting data available for analysis consisted of 120 student ratings (teacher and staff) in reading and 130 student ratings (teacher and staff) in mathematics. Findings indicated a high degree of agreement between teacher and staff ratings. (NQ)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144739

A STUDY OF THE VALIDITY OF TEACHER EVALUATIONS OF STUDENTS IN THE REGION 14 EDUCATION SERVICE CENTER COOPERATIVE FOR MIGRANT EDUCATION

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Region 14/Education Service Center
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS"

Region 14 Education Service Center
Abilene, Texas

March 1976

RC 010087

INTRODUCTION

The evaluation system presently in place in the Region 14 Cooperative for Migrant Education program necessitates that instructional personnel evaluate students on a continuing basis. This is a vital part of the program. Assessments are made using criterion-referenced items that have been developed and validated over the past three years.

In order for the system to function effectively, teachers are required to (1) use the placement tests properly, (2) assess student progress on a continuous basis, and (3) devote the necessary time and effort to accurately assessing of student performance. As a part of the on-going staff development program, teachers in the program have been instructed in the proper use of placement-test items, the importance of continuous assessment of students, and the utilization of results for planning purposes. The extent to which these activities are being performed is the focus of continuous monitoring being performed by the migrant cooperative staff.

PURPOSE OF STUDY

The accuracy with which the teachers in the program are applying the measurement system has until the present time not been studied, although effective program evaluation is dependent upon accurate application by the teachers. This study was conducted to answer the question: How accurately are the teachers in the program assessing their students using the project developed measurement system?

PROCEDURES

The project measurement systems consists of 230 criterion-referenced items designed to measure 230 objectives in reading and mathematics. These range across 7 levels and grades K-6. During the month of February, 1976, a stratified random sample of students (grade level X school district) were selected and matched with a group of randomly selected criterion-referenced items from the project measurement system. The project staff who developed the criterion-referenced measurement system (and on whom inter-rater reliability had been established) visited each school and individually tested the randomly selected students on the randomly selected items. Two measures were noted: the existing teacher rating (taken from student records on file in the classroom) and the results of the staff testing. In both cases, teacher and staff ratings were in terms of whether the student did or did not perform satisfactorily on the objective. In cases where the student was absent on the day the staff member visited the school or where the teacher had not performed an evaluation, the student was omitted from the analysis. The resulting data available for analysis consisted of 120 student ratings (teacher and staff) in the area of reading and 130 student ratings (teacher and staff) in the area of mathematics.

RESULTS

Table 1 presents the results of an analysis of frequency of agreement data. It can be noted that a high degree of agreement was obtained between teacher and staff evaluations of students

Table 1

Agreement Rate Between Teacher and Staff Ratings by Reading and Mathematics Areas

Area	Number of Ratings In Agreement	Out of Agreement	Percent of Agreement
Reading	107	13	89%
Mathematics	112	18	86%

(both better than 85%). Further analysis of the information indicated that the disagreement between rating were almost invariably of the nature where the staff indicated mastery (for the student) and the teacher had indicated on her records non-mastery. One interpretation of this information would be that teachers are not typically able to keep student records updated (a very distinct possibility) thus creating the non-mastery teacher rating versus mastery staff rating pattern. If this is the case, then the reliability of teacher evaluations are even better than indicated by Table 1 data.

Table 2 presents the Phi correlation coefficients results for the teacher and staff ratings for both reading and mathematics. The correlations as expected are positive and relatively high. These data tended to further reflect the substantial agreement between teacher and staff ratings of students.

TABLE 2

Intercorrelation Matrix Depicting Phi Correlation Coefficients for Teacher and Staff Ratings on the Reading and Mathematics Objectives

VARIABLES	Teacher (Reading)	Staff (Reading)	Teacher (Math)	Staff (Math)
Teacher (Reading)	1.00	.78	N/A	N/A
Staff (Reading)	.78	1.00	N/A	N/A
Teacher (Mathematics)	N/A	N/A	1.00	.69
Staff (Mathematics)	N/A	N/A	.69	1.00

SUMMARY AND CONCLUSIONS

This report presents the results of a study conducted at the Region 14 Education Service Center to assess the validity of teacher criterion-referenced assessments performed on students in the migrant education program. The study sought to answer the question: Are teachers in the program able to use the project developed criterion-referenced measurement system to produce valid measures on students?

The results indicated the answer is yes. Substantial agreement was found between teacher ratings and staff ratings. Thus it would appear that the teacher ratings produced as a part of program operations are valid and can be used as input into program evaluation, modification and overall improvement efforts.